N71-28100

NASA CR-118999

AN OPTIMUM SETTLING PROBLEM FOR TIME LAG SYSTEMS*

by

Marc Q. Jacobs and Ti-Jeun Kao
Center for Dynamical Systems
Division of Applied Mathematics
Brown University
Providence, Rhode Island

*This research was supported in part the Air Force Office of Scientific Research under AF-AFOSR 693-67D, and in part by the National Aeronautics and Space Administration under NGL 40-002-015.

AN OPTIMUM SETTLING PROBLEM FOR TIME LAG SYSTEMS

by

Marc Q. Jacobs and Ti-Jeun Kao

§1. Introduction. The problem to be considered is motivated by the following considerations. Suppose a linear control system of the form

$$(1.1) \qquad \dot{x} = A_0 x(t) + A_1 x(t-\sigma) + B_0 u(t), \qquad \sigma > 0$$

is given. Starting from a given initial function φ at time a the problem is to bring the system to rest at a specified time $b > a + \sigma$ in such a way that a certain performance index $J(x,u)$ is minimized on a class U of admissible controls. The rest position or equilibrium position is in this case the zero function on $[-\sigma,0]$ (cf. [8]). Following Hale [8] if $x: [a-\sigma,b] \to R^n$ and $t \in [a,b]$, then we use x_t for the function on $[-\sigma,0]$ defined by $x_t(\theta) = x(t+\theta)$, $-\sigma \le \theta \le 0$. In this notation the boundary conditions for the above problem are $x_a = \varphi$, $x_b = 0$. Our purpose in this note is to present a solution to this type of problem by the classical method of Lagrange multipliers in a Banach space. The actual systems we consider will have the form

$$(1.1') \qquad \dot{x}(t) = f(t,x(t),x(t-\sigma),u(t),u(t-\tau)), \quad \sigma,\tau > 0$$

where f is in general a nonlinear function (cf. Section 2), with boundary conditions

$$(1.2) \qquad x_a = \varphi, \qquad x_b = \psi.$$

The problem is to minimize a given functional

$$(1.3) \qquad J(x,u) = \int_a^b L(t,x,u)dt$$

on a suitable class of admissible pairs (x,u) satisfying (1.1') and (1.2). (If one wants to bring a linear system of the form (1.1') to equilibrium when there are lags in the controls, then appropriate functional end conditions on the admissible controls u should also be imposed as in [3]).

When this problem is formulated as a Lagrange multiplier problem in Sobolev space W_2^1 (cf. [9]) a rather complete solution to the above variational problem is possible. In particular, we obtain necessary and sufficient conditions for the Lagrange multiplier problem to be regular [16, 17] in the form of a controllability result, thereby giving conditions for normality in the sense of the calculus of variations. A set of necessary conditions for an optimal control is derived, and for a special class of convex problems these conditions are also sufficient. Existence and uniqueness theorems for an optimal control are also established.

There are good bibliographies on recent developments for control problems involving functional differential equations in [1], [2], [3]. However, we have not found any literature on the optimization problem proposed above with the exception of [11] and [12]. Kao in [11] considered a special case of the above problem. There the Lagrange multiplier problem was formulated in the B-space C of continuous functions with the norm of uniform convergence. In that formulation conditions for a regular Lagrange multiplier problem could not be obtained. Nonetheless formal calculations led to a valid sufficient condition. In [12] Kent obtained results for a general class of problems (involving neutral functional differential equations) which at least partially cover those obtained herein. There is in [12] the explicit hypothesis that function ψ in (1.2) be continuously differentiable and other than the zero function. However, this condition can easily be relieved by a number of devices. Kent did not obtain conditions for the regularity (or normality) of his problem as we are able to do (cf. Lemma 3.1). In addition, Kent's method of solution (based on Neustadt's work in [18], [19]) differs substantially from the more classical one which we give. The approach we have taken indicates that the Sobolev space W_2^1 is a rather attractive choice of the state space for our control problem.

§2. Terminology and Assumptions. Let $\mathcal{D}$ be an open subset of a Banach space E, and let F be a Banach space. If $H: \mathcal{D} \to F$ is Frechet differentiable at $x_0 \in \mathcal{D}$, then we use $H'(x_0)$ to denote the Frechet differential of H at x_0. Suppose H is continuously Frechet differentiable on $\mathcal{D}$. Then $x_0 \in \mathcal{D}$ is called a regular point of the transformation H if $H'(x_0)$ is a surjection (i.e., maps E onto F). If $E = E_1 \times E_2$, $x_0 = (a_1, a_2) \in \mathcal{D}$, then we will use the subscript notation of Dieudonne's text [5, pg. 167] for partial derivatives. That is, the Frechet derivative of $x_1 \mapsto H(x_1, a_2)$, $(x_1, a_2) \in \mathcal{D}$ (respectively, $x_2 \mapsto H(a_1, x_2)$, $(a_1, x_2) \in \mathcal{D}$) is denoted by $D_1H(a_1, a_2)$ (respectively $D_2H(a_1, a_2)$).

The Banach spaces which are used in this paper are actually real Hilbert spaces. If E is a real Hilbert space we use $\langle x, y \rangle$ for the scalar product of x and $y \in E$. We also use $\|\cdot\|$ for any norms that come into the discussion including the norm of a linear operator. The context will make clear in which space the norm is being applied, and there will be no need for distinguishing subscripts on the various norms. The notation R^p denotes Euclidean space of p dimensions, $p = 1, 2, \ldots$. Vectors $x \in R^p$ will be written as column vectors. We use A^* to denote the adjoint of a given linear operator. Thus in the case of a matrix A, A^* denotes the transposed matrix.

In all situations where the notion of measurability intervenes Lebesgue measure is understood. If $[\alpha, \beta]$ is a compact interval, then we use $L_2([\alpha,\beta],R^p)$ in its usual sense [6] to denote the Hilbert space of all "square integrable" functions on $[\alpha,\beta]$ with functions identified if they are equal almost everywhere (a.e.) on $[\alpha,\beta]$. We use $W_2^1([\alpha,\beta],R^p)$ to denote the collection of all absolutely continuous $Z: [\alpha,\beta] \to R^p$ such that $t \mapsto \frac{dZ(t)}{dt} \equiv \dot{Z}(t)$ is in $L_2([\alpha,\beta]),R^p)$. The inner product on $W_2^1([\alpha,\beta],R^p)$ is defined by

$$(2.1) \qquad \langle Z_1,Z_2\rangle = \langle Z_1(\alpha),Z_2(\alpha)\rangle + \int_\alpha^\beta \langle \dot{Z}_1(t),\dot{Z}_2(t)\rangle dt.$$

The space $W_2^1([\alpha,\beta],R^p)$ with this inner product is a Hilbert space.

Now we discuss the assumptions that will be needed for the functions f and L which were formally introduced in equations (1.1') and (1.3). The functions $f: R \times R^n \times R^n \times R^m \times R^m \to R^n$ and $L: R \times R^n \times R^m \to R$ are assumed to be continuous. Generic points in $R \times R^n \times R^n \times R^m \times R^m$ (respectively, $R \times R^n \times R^m$) are denoted by (t,r_1,r_2,r_3,r_4) (respectively, (t,p_1,p_2)). For each fixed t functions f and L are taken to be continuously differentiable in the remaining arguments. Let $X \equiv W_2^1([a-\sigma,b],R^n)$, $\sigma > 0$ and let

$$(2.2) \qquad U \equiv \{u \in L_2([a-\tau,b],R^m) | u(t) = 0,\ a - \tau \leq t \leq a\}, \quad \tau > 0.$$

Suppose also that $b > \max\ (a+\sigma, a+\tau)$. The set U is called the set of <u>admissible controls</u>, and we will consider U to be identified in the natural way with $L_2([a,b],R^m)$. The problem we propose to solve is

$$\text{(P)} \quad \begin{array}{l} \text{Minimize (locally)} \quad J(x,u) \equiv \int_a^b L(t,x,u)dt \\ \text{with constraints: } x \in X, \quad u \in U \end{array}$$

$$\dot{x}(t) = f(t,x(t),x(t-\sigma),u(t),u(t-\tau)), \text{ a.e. on } [a,b],$$

$$x_a = \varphi, \quad x_b = \psi$$

where φ and ψ are fixed elements of $X_\sigma \equiv W_2^1([-\sigma,0],R^n)$. Other kinds of end conditions, e.g.,

$$G(x_b) = 0$$

where $G: X_\sigma \to R^r$ is continuously differentiable can also be handled essentially by repeating the arguments for our solution to (P). Problem (P) represents the most difficult case since it is here that the regularity conditions will be most difficult to satisfy. At any rate in this paper only problem (P) will be treated. It should also be mentioned that there are various devices in the classical theory of calculus of variations for treating additional constraints on controls $u \in U$, e.g., $|u^i| \leq 1$, $i = 1,2,\ldots,m$ where $u = (u^1,\ldots,u^m)^*$. An instance of this is illustrated in the examples at the end of the paper (cf., also Remark 5.2).

Although we still do not have enough assumptions on the functions f and L for us to set up the Lagrange multiplier problem which we have in mind, nonetheless, let us "define" formally the operators that will be needed and then it will be clear what additional hypotheses are needed: Let $\mathscr{A}: X \times U \to X$ be "defined" by

$$(2.3)\quad \mathscr{A}(x,u)(t) = \begin{cases} \dot{x}(t) - \varphi(t-a), \text{ if } a - \sigma \leq t \leq a \\ x(t) - \varphi(t-a) - \int_a^t f(s,x(s),x(s-\sigma),u(s),u(s-\tau))ds, \\ \text{if } a \leq t \leq b. \end{cases}$$

Thus if $\mathscr{A}$ is to be well defined for each $(x,u) \in X \times U$ some growth conditions are going to have to be imposed on f in the arguments r_3 and r_4. Moreover, we will need $\mathscr{A}$ to be continuously Frechet differentiable. Taking into account the norms on X and U one sees that rather stringent growth restrictions on f are required. What we actually need are hypotheses for L and f that assure:

(2.4) $J: X \times U \to R$ and $\mathscr{A}: X \times U \to X$ are both well defined and continuously differentiable.

The hypotheses we make are: there exist functions $M: R \times R^n \times R^n \to R$, $m: R \times R^n \to R$ which are bounded on bounded sets such that

$$\text{(a)}\quad \|f(t,r_1,r_2,r_3,r_4) - f(t,r_1,r_2,r_3',r_4')\| \leq M(t,r_1,r_2)[\|r_3-r_3'\| + \|r_4-r_4'\|],$$

$$\|D_i f(t,r_1,r_2,r_3,r_4) - D_i f(t,r_1,r_2,r_3',r_4')\| \leq$$

(2.5) $$M(t,r_1,r_2)[\|r_3 - r_3'\| + \|r_4 - r_4'\|],$$

$$i = 1,2, r_1,r_2 \in R^n, \text{ and } r_3,r_4,r_3',r_4' \in R^m$$

$$\|D_i f(t,r_1,r_2,r_3,r_4)\| \leq M(t,r_1,r_2), \ i = 3,4,$$

$$r_1,r_2 \in R^n, \ r_3,r_4 \in R^m;$$

(b) $$\|L(t,p_1,p_2) - L(t,p_1,p_2')\| \leq m(t,p_1)\|p_2 - p_2'\|,$$

$$p_1 \in R^n, \ p_2,p_2' \in R^m,$$

$$\|D_i L(t,p_1,p_2) - D_i L(t,p_1,p_2')\| \leq m(t,p_1)\|p_2 - p_2'\|,$$

$$i = 1,2, \ p_1 \in R^n, \ p_2,p_2' \in R^m.$$

Here $D_i f$ (respectively $D_i L$) denotes the partial derivative with respect to the argument r_i (respectively p_i) with the corresponding subscript.

With these assumptions it can be shown that (2.4) is satisfied. Moreover, if $(\bar{x},\bar{u}) \in X \times U$, then the Frechet differential of $\mathscr{A}$ at $(\bar{x},\bar{u})$ in the direction $(h,v) \in X \times U$ is given by

$$(2.6)\qquad \mathscr{A}'(\bar{x},\bar{u})(h,v)(t) = \begin{cases} h(t), \text{ if } a-\sigma \le t \le a \\ h(t) - \int_a^t [D_1\bar{f}(s)h(s) + D_2\bar{f}(s)h(s-\sigma) \\ + D_3\bar{f}(s)v(s) + D_4\bar{f}(s)v(s-\tau)]ds, \\ \text{if } a \le t \le b, \end{cases}$$

where $D_i\bar{f}(s)$ is defined by

$$(2.7)\qquad D_i\bar{f}(s) = D_if(s,\bar{x}(s),\bar{x}(s-\sigma),\bar{u}(s),\bar{u}(s-\tau)), \qquad a \le s \le b,$$

$i = 1,2,3,4$. A similar[1] meaning is assigned to $D_i\bar{L}(s)$, $i = 1,2$, and we have

$$(2.8)\qquad J'(\bar{x},\bar{u})(h,v) = \int_a^b \langle D_1\bar{L}(s),h(s)\rangle + \langle D_2\bar{L}(s),v(s)\rangle ds,$$

$h,v \in X \times U$. Let $\mathscr{B}: X \times U \to X_\sigma$ be defined by

$$(2.9)\qquad \mathscr{B}(x,u) = x_b - \psi, \qquad (x,u) \in X \times U.$$

The mapping $H \equiv \mathscr{A} \times \mathscr{B}: X \times U \to X \times X_\sigma$ is defined by

$$(2.10)\qquad H(x,u) = (\mathscr{A}(x,u),\ \mathscr{B}(x,u)).$$

---

[1] $D_i\bar{L}(s)$, $i = 1,2$ are treated as column vectors.

The optimization problem (P) has now been transformed into the following Lagrange multiplier problem:

(PL) Minimize (locally) $J(x,u)$ on the manifold

$$M = \{(x,u) \in X \times U \mid H(x,u) = 0\}.$$

Remark 2.1. The unpleasant conditions (2.5) (a), (b) where imposed because we made a choice of $U \subset L_2([a-\tau,b],R^m)$ and required (2.4). Was this merely an unfortunate choice of the function space U? Certainly if we had chosen U to be the corresponding subspace of $L_\infty([a-\tau,b],R^n)$, the result (2.4) could have been obtained directly from the differentiability conditions on f and L. But the investigation of conditions for the regularity of the transformation H in problem (PL) given in Section 3 (to follow) would dictate the choice of the spaces X and X_σ as the corresponding Sobolev spaces W^1_∞. While we do not want to rule out this approach, it seems clear that the solution to (PL) with this change of spaces X, X_σ, and U will be substantially more complicated. Likewise for a number of problems a suitable B-space C of continuous functions (with the norm of uniform convergence) turns out to be a pleasant choice for X, X_σ, and U. However, simple examples of problem (P) show that some discontinuities in the controls u must be allowed. Moreover, if the state spaces X and X_σ are spaces of continuous functions C, then the regularity condition that will be needed on H (see Section 3) will not be true (cf. [11]). In effect this regularity condition on H forces a certain type of "compatibility" between the spaces X, X_σ, and U. We do not claim that some other choices of spaces X, X_σ, U could not also lead to a correct solution to (PL), but it does appear that the more or less complete nature of the solution to problem (PL) we are able to obtain using the Sobolev spaces $W_2^1([a-\sigma,b],R^n)$, $W_2^1([-\sigma,0],R^n)$ for the state spaces X and X_σ respectively and using L_2 controls does offer some fairly strong support for our choice of spaces.

§3. Regularity and Controllability. In this section the question of regularity of the transformation H in (PL) is considered. Only the special problem with no lags in the controls will be treated, i.e., we assume that f is independent of r_4 so that $f\colon R \times R^n \times R^n \times R^m \to R^n$. The case with lags in the control can also be treated, but only at the expense of considerably more technical detail. For notational convenience we adopt the notation

$$(3.1) \qquad \begin{aligned} D_i\bar{f}(t) &= P_i(t), \quad i = 1,2 \\ D_3\bar{f}(t) &= Q_1(t) \end{aligned}$$

where $D_i\bar{f}(t)$ is defined in (2.7). Now $(\bar{x},\bar{u})$ is a regular point of the transformation H in (2.10) means that the bounded linear operator $H'(x,u)\colon X \times U \to X \times X_\sigma$ is surjective. The mapping in (2.9) has Frechet derivative given by

$$(3.2) \qquad \mathscr{L}'(\bar{x},\bar{u})(h,v) = h_b.$$

Taking advantage of (2.6) and (3.2) it is noted that $(\bar{x},\bar{u})$ being a regular point of H is equivalent to the following controllability condition: For each choice of $(y,\lambda) \in X \times X_\sigma$ there is an $(h,v) \in X \times U$ such that

$$(3.3) \qquad \begin{aligned} &\dot{y}(t) = \dot{h}(t) - P_1(t)h(t) - P_2(t)h(t-\sigma) - Q_1(t)v(t), \text{ a.e. on } [a,b], \\ &h_a = y_a, \quad h_b = \lambda. \end{aligned}$$

Extend λ to all of R by defining $\lambda(t) = \lambda(0)$, $t > 0$ and $\lambda(t) = \lambda(-\sigma)$, $t < -\sigma$. Define

$$z(t) = h(t) - \lambda(t-b).$$

Then (3.3) can be rewritten as

$$(3.4)\quad \begin{aligned} \dot{z} &= P_1(t)z(t) + P_2(t)z(t-\sigma) + Q_1(t)v(t) + \zeta(t), \text{ a.e. on } [a,b], \\ z_a &= y_a - \lambda(-\sigma), \qquad z_b = 0, \end{aligned}$$

where

$$(3.5)\quad \zeta(t) = \dot{y}(t) - \dot{\lambda}(t-b) + P_1(t)\lambda(t-b) + P_2(t)\lambda(t-b-\sigma).$$

If $(\bar{x},\bar{u})$ is a regular point of H, then there is a $v \in U$ such that (3.4) is satisfied. Hence

$$(3.6)\quad P_2(t)z(t-\sigma) + Q_1(t)v(t) + \zeta(t) = 0 \text{ a.e. on } [b-\sigma,b].$$

The equation (3.6) must have a solution $v \in L_2([b-\sigma,b],R^m)$ for every admissible choice of ζ of the form (3.5). On the interval $[b-\sigma,b]$ function $t \mapsto \dot{y}(t)$ is an arbitrary L_2 function. It easily follows that the mapping $Q\colon L_2([b-\sigma,b],R^m) \to L_2([b-\sigma,b],R^n)$ defined by

$$(3.7) \qquad (Qv)(t) = Q_1(t)v(t), \quad b - \sigma \leq t \leq b,$$

must be surjective. Using the conditions (2.5)(a) we find that $t \mapsto \|Q_1(t)\|^2$, $a \leq t \leq b$ is integrable so that Q is actually a bounded linear operator. Since Q is surjective it follows [17, pg. 161] that QQ^* is also a surjection where Q^* is the adjoint of Q. One can easily verify that

$$(3.8) \qquad (QQ^*\xi)(t) = Q_1(t)Q_1^*(t)\xi(t), \quad b - \sigma \leq t \leq b$$

for $\xi \in L_2([b-\sigma,b],R^n)$. Choose the natural basis $e_1,\ldots,e_n$ for R^n, and consider these vectors in R^n as vectors in $L_2([b-\sigma,b],R^n)$ in the obvious way. Then there exist $\xi_1,\ldots,\xi_n \in L_2([b-\sigma,b],R^n)$ such that

$$(QQ^*\xi_i)(t) = e_i = Q_1(t)Q_1^*(t)\xi_i(t), \quad \text{a.e. on } [b-\sigma,b].$$

Consequently, for almost every $t \in [b-\sigma,b]$ the $n \times n$ matrix $Q_1(t)Q_1^*(t)$ is invertible. Thus QQ^* is a bijective continuous linear mapping on $L_2([b-\sigma,b],R^n)$. The open mapping theorem [6] assures us that $(QQ^*)^{-1}$ is also a bounded linear mapping. By Holder's inequality and the boundedness of $(QQ^*)^{-1}$ it is determined that

$$\int_{b-\sigma}^{b} \|(Q_1(t)Q_1^*(t))^{-1}\xi(t)\|dt \leq [\sigma \int_{b-\sigma}^{b} \|(Q_1(t)Q_1^*(t))^{-1}\xi(t)\|^2 dt]^{1/2}$$
$$\leq \sigma^{1/2}\|(QQ^*)^{-1}\| \; \|\xi\|$$

for each $\xi \in L_2([b-\sigma,b],R^n)$. Applying the converse to Holder's inequality [21, pg. 277] we obtain that the function

$$(3.9) \qquad t \mapsto \|(Q_1(t)Q_1^*(t))^{-1}\|^2, \qquad b-\sigma \le t \le b$$

is well defined almost everywhere on $[b-\sigma,b]$ and is integrable on $[b-\sigma,b]$.

There is an immediate partial converse to this result. Suppose that $Q_1(t)Q_1^*(t)$ is invertible almost everywhere on $[a,b]$ and suppose the function in (3.9) is integrable on the interval $[b-\sigma,b]$. Let $X(t,s)$ be the fundamental matrix for the homogeneous equation corresponding to (3.4), i.e., $X(t,s)$ is an $n \times n$ matrix such that $s \mapsto X(t,s)$, $s < t$ is the absolutely continuous solution to

$$\frac{\partial X}{\partial s}(t,s) = -X(t,s)P_1(s) - X(t,s+\sigma)P_2(s+\sigma), \quad s < t$$

$$X(t,t) = I = n \times n \text{ identity matrix}$$

$$X(t,s) = 0, \quad s > t$$

[7, pg. 359]. Just as in [22] one can verify that the matrix

$$(3.10) \qquad \int_a^{b-\sigma} X(b-\sigma,s)Q_1(s)Q_1^*(s)X^*(b-\sigma,s)ds$$

has rank n. For suppose the matrix has rank less than n. Then there is a nonzero vector $\eta \in R^n$ such that

$$Q_1^*(s)X^*(b-\sigma,s)\eta = 0 \quad \text{a.e. on} \quad [a,b-\sigma].$$

Since, however, $Q_1^*(s)$ must have rank n a.e. on $[a,b]$ we have that

$$X^*(b-\sigma,s)\eta = 0 \quad \text{a.e. on} \quad [a,b-\sigma]$$

and so $\eta = 0$ contrary to our assumption. Since the matrix (3.10) has rank n one can select a suitable $\eta \in R^n$ such that the function

$$s \mapsto v(s) \equiv Q_1^*(s)X^*(b-\sigma,s)\eta, \quad a \leq s \leq b - \sigma$$

will provide a solution on the interval $[a,b-\sigma]$ to the differential equation in (3.4) satisfying the boundary conditions

$$(3.11) \qquad z_a = y_a - \lambda(-\sigma), \quad z(b-\sigma) = 0.$$

This follows at once from the variation of parameters formula [7, pg. 361] and the fact that the matrix in (3.10) is invertible. Now the assumptions on Q_1 assure us that v can be extended to an L_2 function on $[a,b]$ in such a way as to assure (3.6) (and hence (3.4)) by taking $v(t) = Q_1^*(t)\xi(t)$, $b - \sigma \leq t \leq b$ for an appropriate choice of $\xi \in L_2([b-\sigma,b],R^n)$ (QQ^* has a bounded inverse as above). We have proved the following lemma.

Lemma 3.1. If $(\bar{x},\bar{u}) \in X \times U$ is a regular point of the transformation H in problem (PL), then the matrix $Q_1(t)$ has rank n for almost all $t \in [b-\sigma,b]$ (so that $Q_1(t)Q_1^*(t)$ is invertible for almost all $t \in [b-\sigma,b]$) and $t \mapsto \|(Q_1(t)Q_1^*(t))^{-1}\|^2$ is integrable on $[b-\sigma,b]$. Moreover, if the matrix $Q_1(t)$ has rank n a.e., on $[a,b]$ and if the function in (3.9) is integrable on $[b-\sigma,b]$, then $(\bar{x},\bar{u})$ is a regular point of the transformation H.

Remark 3.1. The necessary condition of the preceding lemma is clearly also sufficient in the case where Q_1 is not dependent on time. On the other hand for sufficiency in the time dependent case one can clearly relax the assumptions on $Q_1(t)$ on the interval $[a,b-\sigma]$ to any assumption assuring that the matrix in (3.10) has rank n thereby guaranteeing the existence of an L_2 control v such that the differential equation in (3.4) has a solution with boundary conditions (3.11) (cf., [22]). The conditions for regularity of H are then the analogs of corresponding results for ordinary control problems [17, pg. 256]. If the right end condition in (1.2) involved a point constraint $x(b) = x_1 \in R^n$ or more generally $G(x_b) = 0$ where $G: X_\sigma \to R^r$ is a "suitable" continuously differentiable function, then conditions for the regularity of the associated Lagrange multiplier problem are easier to obtain. The expected connections with the controllability of the corresponding linear system (3.3) were established in [11]. It is also instructive to compare Lemma 3.1 with the assumptions in [10, pg. 352].

§4. Necessary Conditions. We can now apply the classical Lagrange multiplier theorem [16, pg. 209 or 17, pg. 243] to problem (PL).

Theorem 4.1. Let $(\bar{x},\bar{u})$ be a solution to (PL), and let $(\bar{x},\bar{u})$ be a regular point of the transformation H in (2.10). Then there is a function $\eta: [a,\infty) \to R^n$ such that $\eta|[a,b-\sigma]$ is an absolutely continuous solution to

$$(N_1) \quad \begin{aligned} -\dot{\eta}(t)^* = {} & (D_1L(t,\bar{x}(t),\bar{u}(t)))^* + \eta(t)^* D_1 f(t,\bar{x}(t),\bar{x}(t-\sigma),\bar{u}(t),\bar{u}(t-\tau)) \\ & + \eta(t+\sigma)^* D_2 f(t+\sigma,\bar{x}(t+\sigma),\bar{x}(t),\bar{u}(t+\sigma),\bar{u}(t+\sigma-\tau)) \end{aligned}$$

a.e., on $[a,b-\sigma]$, and there is a $\mu \in X_\sigma$ such that

$$\rho(t) \equiv \eta(t) - \dot{\mu}(t-b), \quad b - \sigma \leq t \leq b$$

is absolutely continuous and satisfies

$$(N_2) \quad -\dot{\rho}(t)^* = (D_1L(t,\bar{x}(t),\bar{u}(t)))^* + \eta(t)^* D_1 f(t,\bar{x}(t),\bar{x}(t-\sigma),\bar{u}(t),\bar{u}(t-\tau))$$

a.e., on $[b-\sigma,b]$ and

$$(N_3) \quad \begin{cases} \eta(b) = \dot{\mu}(0) \\ \eta(b-\sigma) - \eta((b-\sigma)^-) = \dot{\mu}(-\sigma) - \mu(-\sigma) \\ \eta(t) = 0, \quad t > b \end{cases}$$

Moreover, the following equalities are satisfied

$$(N_4)\qquad \begin{aligned} &(D_2L(t,\bar{x}(t),\bar{u}(t)))^* + \eta(t)^* D_3 f(t,\bar{x}(t),\bar{x}(t-\sigma),\bar{u}(t),\bar{u}(t-\tau)) \\ &\quad + \eta(t+\tau)^* D_4 f(t+\tau,\bar{x}(t+\tau),\bar{x}(t-\sigma+\tau),\bar{u}(t+\tau),\bar{u}(t)) \\ &\quad = 0 \quad \text{a.e., on } [a,b-\tau], \text{ and} \end{aligned}$$

$$(N_5)\qquad \begin{aligned} &(D_2L(t,\bar{x}(t),\bar{u}(t)))^* + \eta(t)^* D_3 f(t,\bar{x}(t),\bar{x}(t-\tau),\bar{u}(t),\bar{u}(t-\tau)) = 0 \\ &\quad \text{a.e., on } [b-\tau,b]. \end{aligned}$$

Proof. The Lagrange Multiplier Theorem [16, pg. 209; 17, pg. 243] reveals that there is a $\lambda \in X$ and $\mu \in X_\sigma$ such that

$$(4.1)\qquad J'(\bar{x},\bar{u})(h,v) + \langle \mathscr{A}'(\bar{x},\bar{u})(h,v),\lambda\rangle + \langle h_b,\mu\rangle = 0$$

for every $h \in X$ and $v \in U$. Using the notation in (2.7) the result in (4.1) yields with the aid of (2.6), (2.1) and (2.8) the following two equations

$$(4.2)\qquad \begin{aligned} &\int_a^b [\langle D_1\bar{L}(t),h(t)\rangle + \langle \dot{h}(t) - D_1\bar{f}(t)h(t) - D_2\bar{f}(t)h(t-\sigma),\dot{\lambda}(t)\rangle]dt \\ &\quad + \langle h(b-\sigma),\mu(-\sigma)\rangle + \int_{-\sigma}^0 \langle \dot{h}(b+\theta),\dot{\mu}(\theta)\rangle d\theta = 0 \end{aligned}$$

for every $h \in X$ with $h_a = 0$, and

$$(4.3)\quad \int_a^b [\langle D_2\bar{L}(t),v(t)\rangle - \langle D_3\bar{f}(t)v(t) + D_4\bar{f}(t)v(t-\tau),\dot{\lambda}(t)\rangle]dt = 0$$

for every $v \in U$. Extend the function λ by defining

$$(4.4)\qquad \lambda(t) = \lambda(b),\quad t > b.$$

Taking $h_a = h_b = 0$ in (4.2) one can readily verify

$$(4.5)\quad \begin{aligned} \int_a^{b-\sigma} &[\langle D_1\bar{L}(t),h(t)\rangle - \langle (D_1\bar{f}(t))^* \dot{\lambda}(t) + (D_2\bar{f}(t+\sigma))^* \dot{\lambda}(t+\sigma),h(t)\rangle \\ &+ \langle \dot{\lambda}(t),\dot{h}(t)\rangle]dt = 0. \end{aligned}$$

A modification of the Fundamental Lemma of the Calculus of Variations [10, Lemma 15.2, pg. 51] gives that there is a $c \in R^n$ such that

$$\dot{\lambda}(t) = c + \int_a^t [D_1\bar{L}(s) - (D_1\bar{f}(s))^* \dot{\lambda}(s) - (D_2\bar{f}(s+\sigma))^* \dot{\lambda}(s+\sigma)]ds$$

a.e., on $[a,b-\sigma]$. We may, henceforth, assume that $\dot{\lambda}$ is actually absolutely continuous on $[a,b-\sigma]$ and

$$(4.6)\quad \ddot{\lambda}(t)^* = (D_1\bar{L}(t))^* - \dot{\lambda}(t)^*D_1\bar{f}(t) - \dot{\lambda}(t+\sigma)^*D_2\bar{f}(t+\sigma)$$

a.e., on $[a,b-\sigma]$. Return now to the situation in (4.2), but this time take $h(t) = 0$ on $[a-\sigma,b-\sigma]$ and we arrive at

$$\int_{b-\sigma}^{b} [\langle D_1\overline{L}(s),h(s)\rangle - \langle (D_1\overline{f}(s))^* \dot{\lambda}(s),h(s)\rangle \tag{4.7}$$
$$+ \langle \dot{\lambda}(s) + \dot{\mu}(s-b),\dot{h}(s)\rangle]ds = 0.$$

Once more applying the modified version of the fundamental variational lemma [10, pg. 51] it is determined that if we define

$$\rho(t) \equiv \dot{\lambda}(t) + \dot{\mu}(t-b), \quad b - \sigma \leq t \leq b, \tag{4.8}$$

then there is a vector $K \in R^n$ such that

$$\rho(t) = K + \int_{b-\sigma}^{t} [D_1\overline{L}(s) - (D_1\overline{f}(s))^* \dot{\lambda}(s)]ds$$

a.e., on $[b-\sigma,b]$. Therefore, $\rho(t)$ may be assumed absolutely continuous on $[b-\sigma,b]$ and

$$\dot{\rho}(t)^* = (D_1\overline{L}(t))^* - \dot{\lambda}(t)^* D_1\overline{f}(t), \text{ a.e. on } [b-\sigma,b]. \tag{4.9}$$

This two step procedure leading to (4.6), (4.8), and (4.9) may leave a possible ambiguity in the assignment of a value to $\dot{\lambda}$ at $b - \sigma$. However, from (4.2) we see that

$$\int_{a}^{b-\sigma} [\langle D_1\overline{L}(t),h(t)\rangle - \langle (D_1\overline{f}(t))^* \dot{\lambda}(t) + (D_2\overline{f}(t+\sigma))^* \dot{\lambda}(t+\sigma),h(t)\rangle \tag{4.10}$$

$$+ \langle \dot{\lambda}(t), \dot{h}(t)\rangle]dt + \langle h(b-\sigma), \mu(-\sigma)\rangle$$

$$+ \int_{b-\sigma}^{b} [\langle D_1\bar{L}(t), h(t)\rangle - \langle (D_1\bar{f}(t))^* \dot{\lambda}(t), h(t)\rangle$$

$$+ \langle \rho(t), \dot{h}(t)\rangle]dt = 0$$

for every $h \in X$ with $h_a = 0$. Since ρ is absolutely continuous on $[b-\sigma, b]$ and $\dot{\lambda}$ is absolutely continuous on $[a, b-\sigma]$, the terms

$$\int_{b}^{b-\sigma} \langle \dot{\lambda}(t), \dot{h}(t)\rangle dt, \quad \int_{b-\sigma}^{b} \langle \rho(t), \dot{h}(t)\rangle dt$$

in (4.10) can be integrated by parts to remove the terms involving $\dot{h}$. Then an obvious limiting process leads at once to

$$\text{(4.11)} \qquad \begin{aligned} \rho(b) &= 0 = \dot{\lambda}(b) + \dot{\mu}(0) \\ \rho(b-\sigma) &= \dot{\lambda}((b-\sigma)^-) + \mu(-\sigma). \end{aligned}$$

Finally, turning to the equation (4.3) (valid for every $v \in U$) it is noted that this equation can be written as

$$\text{(4.12)} \qquad \begin{aligned} &\int_{a}^{b-\tau} [\langle D_2\bar{L}(s), v(s)\rangle - \langle (D_3\bar{f}(s))^* \dot{\lambda}(s), v(s)\rangle \\ &\qquad - \langle (D_4 f(s+\tau))^* \dot{\lambda}(s+\tau), v(s)\rangle]ds \\ &+ \int_{b-\tau}^{b} [\langle D_2\bar{L}(s), v(s)\rangle - \langle (D_3\bar{f}(s))^* \dot{\lambda}(s), v(s)\rangle]ds = 0 \end{aligned}$$

for each $v \in U$. It follows at once that

$$(4.13) \qquad (D_2\bar{L}(t))^* - \dot{\lambda}(t)^* D_3\bar{f}(t) - \dot{\lambda}(t+\tau)^* D_4\bar{f}(t+\tau) = 0$$

a.e., on $[a, b-\tau]$ and

$$(4.14) \qquad (D_2\bar{L}(t))^* - \dot{\lambda}(t)^* D_3\bar{f}(t) = 0$$

a.e., on $[b-\tau, b]$. If we change variables by substituting $\eta = -\dot{\lambda}$ in (4.4), (4.6), (4.8), (4.9), (4.11), (4.12), and (4.13), then we get exactly the necessary conditions (N_1) through (N_5) stated in the theorem.

§5. Existence, Uniqueness, and Sufficiency. For linear systems we have the following simple analog of a result in Lee and Markus [15] for ordinary control problems.

Theorem 5.1. Let the system equations in (P) of Section 2 have the form

$$(5.1) \qquad \dot{x}(t) = A_0(t)x(t) + A_1(t)x(t-\sigma) + B_0(t)u(t) + B_1(t)u(t-\tau)$$

a.e., on $[a,b]$ with end conditions

$$(5.2) \qquad x_a = \varphi, \quad x_b = \psi$$

where φ and ψ are fixed functions in X_σ. Let the cost functional J in (P) have the special form

$$J(x,u) = \int_a^b [f^0(x(s),s) + \frac{1}{2}\langle u(s),N(t)u(s)\rangle]ds.$$

Let the mappings[1] $A_i: [a,b] \to \mathcal{L}(R^n,R^n)$, $B_i: [a,b] \to \mathcal{L}(R^m,R^n)$, $i = 0,1$, and $N: [a,b] \to \mathcal{L}(R^m,R^m)$ be continuous. The transformation $N(t)$ is positive semi-definite for each $t \in [a,b]$. Let $f^0: R^n \times R \to R$ be continuous, and let $x \mapsto f^0(x,t)$, $x \in R^n$ be continuously differentiable and convex for each $t \in R$; the convexity condition

[1] $\mathcal{L}(R^p,R^q)$ denotes the vector space of all linear mappings from R^p into R^q with a suitable norm.

means that for each $t \in R$, z, $x \in R^n$ the inequality,

$$(5.3) \qquad \langle \frac{\partial f^0}{\partial x}(x,t), z-x \rangle \leq f^0(z,t) - f^0(x,t),$$

is satisfied. Then the following conclusions hold:

(A) If N is positive definite, then there is at most one solution $(\bar{x},\bar{u})$ to problem (P).

(B) If $f^0 \geq d$ for some constant d, if N is positive definite, and if there is at least one pair $(x,u) \in X \times U$ satisfying (5.1) and (5.2), then there is an optimal solution to problem (P) (globally).

(C) Let $(\bar{x},\bar{u}) \in X \times U$ satisfy (5.1) and (5.2). If functions $\mu \in X_\sigma$, $\eta: [a,\infty) \to R^n$ exist such that η is absolutely continuous on $[a,b-\sigma]$,

$$\rho(t) \equiv \eta(t) - \dot{\mu}(t-b), \quad b - \sigma \leq t \leq b$$

is absolutely continuous, and such that the following conditions are fulfilled:

$$(N_1) \qquad -\dot{\eta}(t) = \frac{\partial f^0}{\partial x}(\bar{x}(t),t) + A_0^*(t)\eta(t) + A_1^*(t+\sigma)\eta(t+\sigma)$$

a.e., on $[a,b-\sigma]$,

$$(N_2) \qquad -\dot{\rho}(t) = \frac{\partial f^0}{\partial x}(\bar{x}(t),t) + A_0^*(t)\eta(t)$$

a.e., on $[b-\sigma,b]$,

$$(N_3) \qquad \begin{cases} \eta(b) = \dot{\mu}(0), \\ \eta(b-\sigma) - \eta((b-\sigma)^-) = \dot{\mu}(-\sigma) - \mu(-\sigma), \\ \eta(t) = 0 \qquad t > b \end{cases}$$

$$(N_4) \qquad N(t)\bar{u}(t) + B_0^*(t)\eta(t) + B_1^*(t+\tau)\eta(t+\tau) = 0$$

a.e., on $[a,b-\tau]$, and

$$(N_5) \qquad N(t)\bar{u}(t) + B_0^*(t)\eta(t) = 0$$

a.e., on $[b-\tau,b]$, then $(\bar{x},\bar{u})$ is a global solution to the problem (P).

Remark 5.1. Note that even when conclusions (A) and (B) are satisfied choice of η and μ may not be uniquely determined by (N_1) through (N_5) and (5.1), (5.2) so some sufficient condition as in (C) is needed.

Proof. Statement (A) is an immediate consequence of standard properties of linear systems and the strict convexity of function J in the variable u. The assumptions in (B) admit the possibility of determining a sequence $(x_n,u_n) \in X \times U$ satisfying (5.1) and (5.2) such that

$$(5.4) \qquad \lim J(x_n,u_n) = \inf \{J(x,u) \mid (x,u) \in X \times U \text{ satisfies (5.1) and (5.2)}\}$$
$$\equiv m.$$

The special form of the integrand for $J(x,u)$ tells us that the sequence $\{u_n\}$ is bounded in L_2. Therefore, we may extract from $\{u_n\}$ a subsequence which we still call $\{u_n\}$ such that u_n is weakly convergent to a $\bar{u} \in U$. One can easily verify that the sequence $v_n: t \mapsto u_n(t-\tau)$ must then converge weakly to $\bar{v}: t \mapsto \bar{u}(t-\tau)$. Hence the variation of parameters formula [7, pg. 361] applied to (5.1) shows at once that responses x_n to the controls u_n must converge pointwise to a function $\bar{x} \in X$ which is the response to control $\bar{u} \in U$. Actually one could show that the convergence $x_n \to \bar{x}$ is uniform on $[a,b]$ [cf. 20], but this stronger result is not needed here. Using Fatou's lemma and the weak lower semi-continuity of the functional $u \mapsto \int_a^b \langle u(t),N(t)u(t)\rangle dt$ (see [16, pg. 123]), it follows at once that

$$J(\bar{x},\bar{u}) \leq \liminf J(x_n,u_n) = m$$

and $\bar{x},\bar{u}$ satisfy (5.1) and (5.2). Therefore, $J(\bar{x},\bar{u}) = m$. This proves (B).

Let $\bar{x},\bar{u}$ satisfying the assumptions of part (C) be called respectively an <u>extremal response</u> and an <u>extremal control</u>. Define a function $\mathscr{H}: R^m \times [a,b] \to R$ by

$$(5.5) \qquad \mathscr{H}(u,t) = -\frac{1}{2}\langle u,N(t)u\rangle - \langle B_0^*(t)\eta(t) + B_1^*(t+\tau)\eta(t+\tau),u\rangle .$$

Now $\partial^2\mathscr{H}/\partial^2 u = -N(t)$ and $\partial\mathscr{H}(\bar{u}(t),t)/\partial u = 0$ a.e., on $[a,b]$ (note that $\eta(t) = 0$ if $t > b$). Hence we have

$$(5.6) \qquad \int_a^b \mathscr{H}(\bar{u}(t),t)dt \geq \int_a^b \mathscr{H}(u(t),t)dt$$

for every $u \in U$. With a little juggling one can show that

$$(5.7) \qquad \begin{aligned} \int_a^b \mathscr{H}(u(t),t)dt = \int_a^b [&-\frac{1}{2}\langle u(t),N(t)u(t)\rangle - \langle \eta(t),B_0(t)u(t)\rangle \\ &- \langle \eta(t),B_1(t)u(t-\tau)\rangle]dt \end{aligned}$$

for $u \in U$ (recall that $u(t) = 0$, $a - \tau \leq t \leq a$). Define $x^0(t)$, $a \leq t \leq b$ to be the absolutely continuous solution to

$$(5.8) \qquad \begin{aligned} &\dot{x}^0(t) = f^0(x(t),t) + \frac{1}{2}\langle u(t),N(t)u(t)\rangle, \quad \text{a.e., on } [a,b] \\ &x^0(a) = 0 \end{aligned}$$

where $(x,u) \in X \times U$. Suppose $(x,u) \in X \times U$ satisfy (5.1) and (5.2), and suppose $\bar{x}^0(t)$ denotes the solution of (5.8) when $(\bar{x},\bar{u})$ is substituted for (x,u) on the right hand side of (5.8). We want to prove that

$$\bar{x}^0(b) \leq x^0(b). \tag{5.9}$$

In order to do this differentiate $-x^0(t) - \langle \eta(t), x(t) \rangle$ and integrate the result over the interval $[a, b-\sigma]$; then differentiate $-x^0(t) - \langle \rho(t), x(t) \rangle$ and integrate the result over the interval $[b-\sigma, b]$. These computations will result in the following two equations:

$$\begin{aligned} -x^0(b-\sigma) &- \langle \eta((b-\sigma)^-), x(b-\sigma) \rangle + \langle \eta(a), x(a) \rangle = \\ &\int_a^{b-\sigma} [-f^0(x(t),t) - \tfrac{1}{2} \langle u(t), N(t)u(t) \rangle + \langle \frac{\partial f^0}{\partial x} (\bar{x}(t),t), x(t) \rangle \\ &+ \langle A_1^*(t+\sigma)\eta(t+\sigma), x(t) \rangle - \langle \eta(t), A_1(t)x(t-\sigma) \rangle \\ &- \langle \eta(t), B_0(t)u(t) \rangle - \langle \eta(t), B_1(t)u(t-\tau) \rangle]dt, \end{aligned} \tag{5.10}$$

$$\begin{aligned} -x^0(b) &+ x^0(b-\sigma) - \langle \rho(b), x(b) \rangle + \langle \rho(b-\sigma), x(b-\sigma) \rangle = \\ &\int_{b-\sigma}^{b} [-f^0(x(t),t) - \tfrac{1}{2} \langle u(t), N(t)u(t) \rangle + \langle \frac{\partial f^0}{\partial x} (\bar{x}(t),t), x(t) \rangle \\ &- \langle \eta(t), A_1(t)x(t-\sigma) \rangle - \langle \eta(t), B_0(t)u(t) \rangle - \langle \eta(t), B_1(t)u(t-\tau) \rangle \\ &+ \langle \dot{\mu}(t-b), \dot{x}(t) \rangle]dt. \end{aligned} \tag{5.11}$$

Specialize the equations (5.10) and (5.11) to control $\bar{u}$ with response $\bar{x}$, and then solve these equations for $\bar{x}^0(b) - x^0(b)$ (the calculation is somewhat tedious and it is important to use the end conditions (5.2) satisfied by both x and $\bar{x}$ in order to get the cancellation of appropriate terms) to obtain

$$\begin{aligned}\bar{x}_0(b) - x_0(b) = \int_a^b [&-f^0(x(t),t) + f^0(\bar{x}(t),t) - \tfrac{1}{2}\langle u(t), N(t)u(t)\rangle \\ &+ \tfrac{1}{2}\langle \bar{u}(t), N(t)\bar{u}(t)\rangle + \langle \tfrac{\partial f^0}{\partial x}(\bar{x}(t),t), x(t) - \bar{x}(t)\rangle \\ &- \langle \eta(t), B_0(t)(u(t) - \bar{u}(t))\rangle - \langle \eta(t), B_1(t)(u(t-\tau) - \bar{u}(t-\tau))\rangle]dt.\end{aligned} \tag{5.12}$$

Equation (5.12) with the aid of (5.7) can be transformed into the following equation,

$$\begin{aligned}\bar{x}^0(b) - x^0(b) = \int_a^b [&f^0(\bar{x}(t),t) - f^0(x(t),t) + \\ &\langle \tfrac{\partial f^0}{\partial x}(\bar{x}(t),t), x(t) - \bar{x}(t)\rangle + \mathscr{H}(u(t),t) - \mathscr{H}(\bar{u}(t),t)]dt.\end{aligned} \tag{5.13}$$

The convexity assumption on f^0 in (5.3) and (5.6) now give (5.9) which is the desired result.

Remark 5.2. It is clear from the proof of part (C) that inequality (5.6) and the identity (5.7) play the crucial role in the sufficiency condition. Indeed, (N_4) and (N_5) can be replaced by (5.6) and (5.7), and we can then treat a more generous class of problems with constraints on the controls. Moreover, in this connection let functions $h^0: R^m \times R \to R$ and $h: R^m \times R \to R^n$ be introduced with

$$L(t,p_1,p_2) \equiv f^0(p_1,t) + h^0(p_2,t)$$

$$f(t,r_1,r_2,r_3,r_4) \equiv A_0(t)r_1 + A_1(t)r_2 + h(r_3,t)$$

in problem (P) subject to the smoothness hypotheses of Section 2 and growth conditions (2.5). Then define

$$\mathscr{K}(u,t) = -h^0(u,t) - \langle \eta(t), h(u,t) \rangle$$

and in Theorem 5.1 (C) replace (N_4) and (N_5) with inequality (5.6) (of course $-\frac{1}{2}\langle u, N(t)u\rangle$ is replaced by $h^0(u,t)$). With these provisions one again obtains the validity of sufficient condition Theorem 5.1 (C) (cf. [15, pg. 341] and Example 5.2 below).

Example 5.1. Consider the scalar system

$$\begin{aligned} \dot{x} &= -x(t-1) + u(t), \qquad 0 \le t \le 2 \\ x_0 &= 1, \qquad x_2 = 0, \end{aligned} \tag{5.14}$$

with cost functional

$$J(x,u) = \int_0^2 u^2 dt \tag{5.15}$$

and control set $U = L_2[0,2]$. Problem (P) for this example was solved in [11] for an arbitrary initial function $\varphi \equiv x_0$. Here we consider only the special case (5.14). First we point out that the solution to this problem involves more than merely finding $\bar{x},\bar{u}$ which minimize $\int_0^1 u^2 dt$, $u \in U$ subject to (5.14), but with the right end condition replaced by $x(1) = 0$. (The zero level can then be maintained on $[1,2]$ by using $u(t) = \bar{x}(t-1)$ on $[1,2]$.) This procedure yields an incorrect solution to the problem. However, this problem can be solved by standard methods [1] by transforming the original problem into the following equivalent problem:

$$\text{Minimize} \quad J(x,u) = \int_0^1 x^2 + u^2 dt$$

subject to constraints

$$\dot{x}(t) = -x(t-1) + u(t), \qquad 0 \leq t \leq 1$$
$$x_0 = 1, \qquad x(1) = 0.$$

This simple problem was chosen so that it could be solved by two different methods and permit us to check the validity of Theorem 4.1. We present only the solution to problem (P) for (5.14) and (5.15) using the results of this paper. Applying Theorem (5.1) (A), (B) (it is not difficult to show there is at least one admissible pair $(x,u) \in X \times U$ satisfying (5.14)). We see there is a unique optimal solution $(\bar{x},\bar{u})$ to problem (P) for (5.14) and (5.15). Using Theorem 4.1 we get

$$\bar{u}(t) = -\eta(t)/2, \qquad 0 \leq t \leq 2$$
$$\dot{\eta}(t) = \eta(t+1),$$
$$\dot{\rho}(t) = \eta(t) - \dot{\mu}(t-2) \equiv 0, \qquad 1 \leq t \leq 2,$$

so that

$$\eta(t) = \begin{cases} \eta(1) - \eta(0) + \mu(t-1), & 0 \leq t < 1 \\ \dot{\mu}(t-2) & 1 \leq t \leq 2. \end{cases}$$

Therefore

$$(5.16) \qquad \bar{u}(t) = \begin{cases} K + \int_1^t \gamma(s+1)ds & 0 \leq t < 1 \\ \gamma(s) & 1 \leq t \leq 2 \end{cases}$$

where $K = \frac{\eta(0) - \eta(1) - \mu(0)}{2}$, $\gamma(t) = \frac{-\dot{\mu}(t-2)}{2}$, $1 \leq t \leq 2$. Now solve the differential equation (5.14) with the two boundary conditions. One finds that for $1 \leq t \leq 2$

$$(5.17) \qquad \int_1^t \gamma(s)ds - \int_1^t \int_0^{p-1} \int_1^r \gamma(s+1)dsdrdp =$$
$$-K - \int_0^1 \int_1^r \gamma(s+1)dsdr + (t-1) + [-1+K]\,\frac{(t-1)^2}{2}.$$

The complicated integral equation actually reduces to an ordinary differential equation

$$\ddot{\gamma}(t) - \gamma(t) = 0$$

so that $\gamma(t) = ae^{t} + be^{-t}, \quad 1 \leq t \leq 2$. This result is substituted back in (5.17) to get

$$a = \frac{e^{-1}}{e^{2} - 1}$$

$$b = \frac{e^{3}}{e^{2} - 1}$$

$$K = \frac{-2e}{e^{2} - 1} + 1.$$

Using (5.16) we get

$$\bar{u}(t) = \begin{cases} \dfrac{e^{t}}{1 - e^{2}} + \dfrac{e^{2}}{1 - e^{2}} e^{-t} + 1, & 0 \leq t < 1 \\[2ex] \dfrac{e^{t-1}}{1 - e^{2}} + \dfrac{e^{2}}{e^{2} - 1} e^{-(t-1)}, & 1 < t \leq 2 \end{cases}$$

and solving (5.14) with this $\bar{u}$ we find

$$\bar{x}(t) = \begin{cases} \dfrac{e^{t}}{1 - e^{2}} + \dfrac{e^{2}}{e^{2} - 1} e^{-t}, & 0 \leq t \leq 1 \\[2ex] 0 & 1 \leq t \leq 2 \end{cases}$$

Theorem 5.1 (C) applies to give that $(\bar{x},\bar{u})$ above is indeed the optimal solution to our problem.

Example 5.2. This example is also of the simplest type, and is aimed at illustrating how a problem with an additional control constraint $|u| \leq 1$ can be dispatched by a standard device. Take the scalar system

$$(5.18)\qquad \begin{aligned} &\dot{x} = -x(t-1) + \sin v(t), \qquad 0 \leq t \leq 2 \\ &x_0 = 0, \qquad x_2 = \psi, \end{aligned}$$

with cost functional

$$(5.19)\qquad J(x,u) = \int_0^2 \frac{\sin^2 v}{2}\, dt.$$

Suppose there is an optimum solution $(\bar{x},\bar{v})$ to problem (P) with (5.18) and (5.19), then Theorem 4.1 says that

$$(5.20)\qquad \begin{aligned} &\dot{\eta}(t) = \eta(t+1), \qquad 0 \leq t < 1 \\ &\eta(t) = \dot{\mu}(t-2), \qquad 1 \leq t \leq 2 \\ &\cos \bar{v}(t)(\sin \bar{v}(t) + \eta(t)) = 0, \qquad 0 \leq t \leq 2. \end{aligned}$$

With the sufficient condition of Remark 5.2 in view it is seen that choice of $\bar{v}$ so that

$$\sin \bar{v}(t) = \begin{cases} +1, & \eta(t) < -1 \\ -\eta(t), & -1 < \eta(t) < 1 \\ -1, & \eta(t) > 1 \end{cases} \tag{5.21}$$

will give the validity of inequality (5.6) in the modified form of Remark 5.2. Thus it is now only a question of solving the complicated integral equation that results from meeting the boundary conditions (if possible) of (5.18) subject to (5.20) and (5.21). Specializing function ψ of (5.18) so that the right end condition is

$$x(t) = e^{-2} - 2e^{-1} + e^{-1}t, \quad 1 \leq t \leq 2$$

the optimum solution $\bar{x}, \bar{u} = \sin \bar{v}$ is given by

$$\bar{u}(t) = \sin \bar{v}(t) = \begin{cases} -e^{-(t+1)} & , \quad 0 \leq t < 1 \\ e^{-t} & , \quad 1 \leq t \leq 2, \end{cases}$$

$$\bar{x}(t) = \begin{cases} -e^{-(t+1)} - e^{-1}, & 0 \leq t \leq 1 \\ e^{-2} - 2e^{-1} + e^{-1}t, & 1 \leq t \leq 2. \end{cases}$$

Kent [12] solved a number of other examples of a more complicated nature. Application of our results or Kent's (where the two approaches speak to the same class of problems) have yielded identical solutions to example problems.

Acknowledgement: The authors are very grateful to Professors H. T. Banks and J. K. Hale for numerous helpful discussions.

References

1. H.T. Banks, Variational problems involving functional differential equations, SIAM J. Control, 7(1969), 1-17.

2. H.T. Banks and M.Q. Jacobs, The optimization of trajectories of linear functional differential equations, SIAM J. Control, 8(1970), 461-488.

3. H.T. Banks, M.Q. Jacobs, and M.R. Latina, The synthesis of time optimal controls for linear problems with retarded controls, to appear in Journal of Optimization Theory and Applications.

4. D.H. Chyung and E.B. Lee, Linear optimal systems with time delay, SIAM J. Control, 4(1966), 548-575.

5. J. Dieudonné, Foundations of Modern Analysis, Academic Press, New York, 1960.

6. N. Dunford and J.T. Schwartz, Linear Operators, Part I, Interscience, New York, 1966.

7. A. Halanay, Differential Equations, Academic Press, New York, 1966.

8. J. Hale, Functional Differential Equations, Lecture Notes at the University of California, Los Angeles, 1968-1969.

9. D. Henry, Linear autonomous functional differential equations of neutral type in Sobolev space W_2^1, to appear.

10. M.R. Hestenes, Calculus of Variations and Optimal Control Theory, John Wiley and Sons, New York, 1966.

11. T.J. Kao, Optimal Control for Functional Differential Equations, Masters Thesis, Brown University, 1971.

12. G.A. Kent, Optimal Control of Functional Differential Equations of Neutral Type, Doctoral Dissertation, Brown University, 1971.

13. F.M. Kirillova, Applications of functional analysis to the theory of optimal processes, SIAM J. Control, 25-50.

14. F.M. Kirillova and S.V. Čurakov, On the controllability of linear systems with aftereffect, (Russian), Diff. Urav. 3(1967), 436-445.

15. E.B. Lee and L. Markus, Foundations of Optimal Control Theory, John Wiley and Sons, New York, 1967.

16. L.A. Liusternik and V.J. Sobolev, Elements of Functional Analysis, Translation, F. Ungar, New York, 1961.

17. D.G. Luenberger, Optimization by Vector Space Methods, John Wiley and Sons, New York, 1969.

18. L.W. Neustadt, An abstract variational theory with applications to a broad class of optimization problems, II, SIAM J. Control, 5(1967), 90-137.

19. L.W. Neustadt, A general theory of extremals, J. Comput. System Sci., 3(1969), 57-92.

20. W.T. Reid, Some limit theorems for ordinary differential systems, J. Diff. Eqs., 3(1967), 423-439.

21. A.E. Taylor, General Theory of Functions and Integration, Blaisdell, New York, 1965.

22. L. Weiss, Lectures on Controllability and Observability, Technical Note BN-590, University of Maryland, 1969.